

RECEIVED

JUL 02 2003

TECH CENTER

X034



1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/961,083A

DATE: 06/30/2003

TIME: 11:37:26

Input Set : A:\340-SL-12-Dec-02.txt

Output Set: N:\CRF4\06272003\H961083A.raw

3 <110> APPLICANT: Choi et. al.
 5 <120> TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 7 <130> FILE REFERENCE: PB340P2
 9 <140> CURRENT APPLICATION NUMBER: 08/961,083A
 10 <141> CURRENT FILING DATE: 1997-10-30
 12 <150> PRIOR APPLICATION NUMBER: 60/029,960
 13 <151> PRIOR FILING DATE: 1996-10-31
 15 <160> NUMBER OF SEQ ID NOS: 4
 17 <170> SOFTWARE: PatentIn version 3.0
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 2389
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Streptococcus pneumoniae
 24 <220> FEATURE:
 25 <221> NAME/KEY: SITE
 26 <222> LOCATION: (1368)..(1368)
 27 <223> OTHER INFORMATION: n equals a, c, g, or t
 30 <400> SEQUENCE: 1
 31 ttcttacgag ttgggactgt atcaagctag aacggttaag gaaaataatc gtgtttcccta 60
 33 tatagatgga aaacaagcga cgcaaaaaac ggagaatttg actcctgatg agtttagcaa 120
 35 gcgtgaagga atcaatgctg agcaaattcg catcaagata acagaccaag gctatgtcac 180
 37 ttccatggc gaccactatc attattacaa tggtaagggt ctttatgacg ctatcatcag 240
 39 tgaagaatta ctcatgaaag atccaaacta taagctaaaa gatgaggata ttgttaatga 300
 41 ggtcaagggt ggatatgtta tcaaggtaga tggaaaatac tatgtttacc ttaaggatgc 360
 43 tgcccacgcg gataacgtcc gtacaaaaga ggaatcaat cgacaaaaac aagagcatag 420
 45 tcaacatcgt gaagggtgaa ctccaagaaa cgatggtgct gttgccttgg cacgttcgca 480
 47 aggacgctat actacagatg atggttatat cttaatgct tctgatatca tagaggatac 540
 49 tggatgtc tatatcggtc ctcatggaga tcattaccat tacattcccta agaatgagtt 600
 51 atcagctagc gagttggctg ctgcagaagc cttccatatct ggtcgaggaa atctgtcaaa 660
 53 ttaagaacc tatcgccgac aaaatagcga taacacttca agaacaaact gggtaccttc 720
 55 tggaaatcc caagaaacta caaataactaa cacaagcaac aacagcaaca ctaacagtca 780
 57 agcaagtcaa agtaatgaca ttgatagtct cttgaaacag ctctacaaac tgccttgag 840
 59 tcaacgacat gtagaatctg atggccttgt ctttgcattca gcacaaatca caagtcgaac 900
 61 agcttagaggt gttgcagtgc cacacggaga tcattaccat ttcattccctt actctcaat 960
 63 gtcgtattt gaaagaaacgaa tcgctcgat tattccctt cgttgcattt cttttttttt 1020
 65 ggtaccagat tcaaggccag aacaaccaag tccacaaccc actccggaaac ctgtccagg 1080
 67 cccgcaacct gcaaaaaatc taaaataga ctcaattct tctttggta gtcagctgg 1140
 69 agaaaaagggt ggggaaggat atgtattcga agaaaaggcc atctctcggt atgtttgc 1200
 71 gaaagattt ccatctgaaa ctgttaaaaa tcttgcggc aagttatcaa aacaagagag 1260
 73 tggttcacac actttaactg ctaaaaaaaa aaatgttgct cctcgtagcc aagaattttt 1320
 W--> 75 tgataaagca tataatctgt taactgaggc tcataaagcc ttgtttgnaa ataagggtcg 1380
 77 taattctgtat ttccaagcct tagacaattt attagaacgc ttgaatgtat aatcgactaa 1440
 79 taaagaaaaaa ttggtagatg atttattggc attcctagca ccaattaccc atccagagcg 1500

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/961,083A

DATE: 06/30/2003
TIME: 11:37:26

Input Set : A:\340-SL-12-Dec-02.txt
Output Set: N:\CRF4\06272003\H961083A.raw

81	acttggcaaa	ccaaattctc	aaattgagta	tactgaagac	gaagttcgta	ttgctcaatt	1560										
83	agctgataag	tatacacgt	cagatggta	cattttgat	gaacatgata	taatcagtga	1620										
85	tgaaggagat	gcatatgtaa	cgcctcatat	gggcctatgt	cactggattg	aaaaagatag	1680										
87	cctttctgat	aaggaaaaag	ttgcagctca	agcctatact	aaagaaaaag	gtatcctacc	1740										
89	tccatctcca	gacgcagatg	ttaaagcaaa	tccaaactgga	gatagtgcag	cagttattha	1800										
91	caatcggtg	aaagggaaa	aacgaattcc	actcggtcga	cttccatata	tgggtgagca	1860										
93	tacagtttag	gttaaaaacg	gtatattgtat	tattcctcat	aaggatcatt	accataatat	1920										
95	taaatttgct	tggtttgatg	atcacacata	caaagctcca	aatggctata	ccttggaga	1980										
97	tttggggcg	acgattaagt	actacgtaga	acaccctgac	gaacgtccac	attctaata	2040										
99	tggatggggc	aatgccagt	agcatgtgtt	aggcaagaaa	gaccacagt	aagatccaaa	2100										
101	taagaacttc	aaagcggatg	aagagccagt	agagggaaaca	cctgctgagc	cagaagtccc	2160										
103	tcaagtagag	actgaaaaag	tagaagccca	actcaaagaa	gcagaagttt	tgcttgcgaa	2220										
105	agtaacggat	tctagtctga	aagccaatgc	aacagaaact	ctagctggtt	tacgaaataa	2280										
107	tttactctt	caaatttatgg	ataacaata	tatcatggca	gaagcagaaa	aattacttgc	2340										
109	gttgttaaaa	ggaagtaatc	cttcatctgt	aagtaaggaa	aaaataaaac		2389										
111	<210>	SEQ ID NO:	2														
112	<211>	LENGTH:	796														
113	<212>	TYPE:	PRT														
114	<213>	ORGANISM:	Streptococcus pneumoniae														
116	<220>	FEATURE:															
117	<221>	NAME/KEY:	SITE														
118	<222>	LOCATION:	(456)..(456)														
119	<223>	OTHER INFORMATION:	Xaa equals any naturally occurring amino acid														
122	<400>	SEQUENCE:	2														
124	Ser	Tyr	Glu	Leu	Gly	Leu	Tyr	Gln	Ala	Arg	Thr	Val	Lys	Glu	Asn	Asn	
125	1				5				10					15			
127	Arg	Val	Ser	Tyr	Ile	Asp	Gly	Lys	Gln	Ala	Thr	Gln	Lys	Thr	Glu	Asn	
128					20				25					30			
130	Leu	Thr	Pro	Asp	Glu	Val	Ser	Lys	Arg	Glu	Gly	Ile	Asn	Ala	Glu	Gln	
131					35				40					45			
133	Ile	Val	Ile	Lys	Ile	Thr	Asp	Gln	Gly	Tyr	Val	Thr	Ser	His	Gly	Asp	
134					50				55					60			
136	His	Tyr	His	Tyr	Tyr	Asn	Gly	Lys	Val	Pro	Tyr	Asp	Ala	Ile	Ile	Ser	
137					65				70					75			80
139	Glu	Glu	Leu	Leu	Met	Lys	Asp	Pro	Asn	Tyr	Lys	Leu	Lys	Asp	Glu	Asp	
140					85				90					95			
142	Ile	Val	Asn	Glu	Val	Lys	Gly	Gly	Tyr	Val	Ile	Lys	Val	Asp	Gly	Lys	
143					100				105					110			
145	Tyr	Tyr	Val	Tyr	Leu	Lys	Asp	Ala	Ala	His	Ala	Asp	Asn	Val	Arg	Thr	
146					115				120					125			
148	Lys	Glu	Glu	Ile	Asn	Arg	Gln	Lys	Gln	Glu	His	Ser	Gln	His	Arg	Glu	
149					130				135					140			
151	Gly	Gly	Thr	Pro	Arg	Asn	Asp	Gly	Ala	Val	Ala	Leu	Ala	Arg	Ser	Gln	
152					145				150					155			160
154	Gly	Arg	Tyr	Thr	Thr	Asp	Asp	Gly	Tyr	Ile	Phe	Asn	Ala	Ser	Asp	Ile	
155					165				170					175			
157	Ile	Glu	Asp	Thr	Gly	Asp	Ala	Tyr	Ile	Val	Pro	His	Gly	Asp	His	Tyr	
158					180				185					190			
160	His	Tyr	Ile	Pro	Lys	Asn	Glu	Leu	Ser	Ala	Ser	Glu	Leu	Ala	Ala	Ala	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/961,083A

DATE: 06/30/2003
TIME: 11:37:26

Input Set : A:\340-SL-12-Dec-02.txt
Output Set: N:\CRF4\06272003\H961083A.raw

161	195	200	205
163	Glu Ala Phe Leu Ser Gly Arg Gly Asn Leu Ser Asn Ser Arg Thr Tyr		
164	210	215	220
166	Arg Arg Gln Asn Ser Asp Asn Thr Ser Arg Thr Asn Trp Val Pro Ser		
167	225	230	235
169	240		
170	Val Ser Asn Pro Gly Thr Thr Asn Thr Asn Ser Asn Asn Ser Asn		
172	245	250	255
173	Thr Asn Ser Gln Ala Ser Gln Ser Asn Asp Ile Asp Ser Leu Leu Lys		
175	260	265	270
176	Gln Leu Tyr Lys Leu Pro Leu Ser Gln Arg His Val Glu Ser Asp Gly		
178	275	280	285
179	Leu Val Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr Ala Arg Gly Val		
181	290	295	300
182	Ala Val Pro His Gly Asp His Tyr His Phe Ile Pro Tyr Ser Gln Met		
184	305	310	315
185	320		
187	Ser Glu Leu Glu Glu Arg Ile Ala Arg Ile Ile Pro Leu Arg Tyr Arg		
188	325	330	335
189	Ser Asn His Trp Val Pro Asp Ser Arg Pro Glu Gln Pro Ser Pro Gln		
190	340	345	350
191	Pro Thr Pro Glu Pro Ser Pro Gly Pro Gln Pro Ala Pro Asn Leu Lys		
193	355	360	365
194	Ile Asp Ser Asn Ser Ser Leu Val Ser Gln Leu Val Arg Lys Val Gly		
196	370	375	380
197	Glu Gly Tyr Val Phe Glu Glu Lys Gly Ile Ser Arg Tyr Val Phe Ala		
199	385	390	395
200	400		
202	Lys Asp Leu Pro Ser Glu Thr Val Lys Asn Leu Glu Ser Lys Leu Ser		
203	405	410	415
205	420	425	430
206	Ala Pro Arg Asp Gln Glu Phe Tyr Asp Lys Ala Tyr Asn Leu Leu Thr		
208	435	440	445
W-->	208 Glu Ala His Lys Ala Leu Phe Xaa Asn Lys Gly Arg Asn Ser Asp Phe		
209	450	455	460
211	Gln Ala Leu Asp Lys Leu Leu Glu Arg Leu Asn Asp Glu Ser Thr Asn		
212	465	470	475
214	480		
215	Lys Glu Lys Leu Val Asp Asp Leu Leu Ala Phe Leu Ala Pro Ile Thr		
217	485	490	495
218	His Pro Glu Arg Leu Gly Lys Pro Asn Ser Gln Ile Glu Tyr Thr Glu		
220	500	505	510
221	Asp Glu Val Arg Ile Ala Gln Leu Ala Asp Lys Tyr Thr Thr Ser Asp		
223	515	520	525
224	Gly Tyr Ile Phe Asp Glu His Asp Ile Ile Ser Asp Glu Gly Asp Ala		
226	530	535	540
227	Tyr Val Thr Pro His Met Gly His Ser His Trp Ile Gly Lys Asp Ser		
229	545	550	555
230	Leu Ser Asp Lys Glu Lys Val Ala Ala Gln Ala Tyr Thr Lys Glu Lys		
232	565	570	575
233	Gly Ile Leu Pro Pro Ser Pro Asp Ala Asp Val Lys Ala Asn Pro Thr		
	580	585	590

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/961,083A

DATE: 06/30/2003
TIME: 11:37:26

Input Set : A:\340-SL-12-Dec-02.txt
Output Set: N:\CRF4\06272003\H961083A.raw

235	Gly	Asp	Ser	Ala	Ala	Ile	Tyr	Asn	Arg	Val	Lys	Gly	Glu	Lys	Arg	
236	595					600					605					
238	Ile	Pro	Leu	Val	Arg	Leu	Pro	Tyr	Met	Val	Glu	His	Thr	Val	Glu	Val
239	610					615					620					
241	Lys	Asn	Gly	Asn	Leu	Ile	Ile	Pro	His	Lys	Asp	His	Tyr	His	Asn	Ile
242	625					630				635					640	
244	Lys	Phe	Ala	Trp	Phe	Asp	Asp	His	Thr	Tyr	Lys	Ala	Pro	Asn	Gly	Tyr
245		645							645	650					655	
247	Thr	Leu	Glu	Asp	Leu	Phe	Ala	Thr	Ile	Lys	Tyr	Tyr	Val	Glu	His	Pro
248		660							660	665				670		
250	Asp	Glu	Arg	Pro	His	Ser	Asn	Asp	Gly	Trp	Gly	Asn	Ala	Ser	Glu	His
251		675						675	680					685		
253	Val	Leu	Gly	Lys	Lys	Asp	His	Ser	Glu	Asp	Pro	Asn	Lys	Asn	Phe	Lys
254		690				695			695	700						
256	Ala	Asp	Glu	Glu	Pro	Val	Glu	Glu	Thr	Pro	Ala	Glu	Pro	Glu	Val	Pro
257	705					710			710	715					720	
259	Gln	Val	Glu	Thr	Glu	Lys	Val	Glu	Ala	Gln	Leu	Lys	Glu	Ala	Glu	Val
260		725						725		730				735		
262	Leu	Leu	Ala	Lys	Val	Thr	Asp	Ser	Ser	Leu	Lys	Ala	Asn	Ala	Thr	Glu
263		740				740			740	745				750		
265	Thr	Leu	Ala	Gly	Leu	Arg	Asn	Asn	Leu	Thr	Leu	Gln	Ile	Met	Asp	Asn
266		755				755			755	760				765		
268	Asn	Ser	Ile	Met	Ala	Glu	Ala	Glu	Lys	Leu	Leu	Ala	Leu	Leu	Lys	Gly
269		770				770			770	775				780		
271	Ser	Asn	Pro	Ser	Ser	Val	Ser	Lys	Glu	Lys	Ile	Asn				
272	785					790			790	795						
274	<210>	SEQ ID NO:	3													
275	<211>	LENGTH:	37													
276	<212>	TYPE:	DNA													
277	<213>	ORGANISM:	Artificial Sequence													
279	<220>	FEATURE:														
280	<223>	OTHER INFORMATION:	PCR primer													
283	<400>	SEQUENCE:	3													
284	agtccggatcc	ttcttacgag	ttgggactgt	atcaagg											37	
287	<210>	SEQ ID NO:	4													
288	<211>	LENGTH:	40													
289	<212>	TYPE:	DNA													
290	<213>	ORGANISM:	Artificial Sequence													
292	<220>	FEATURE:														
293	<223>	OTHER INFORMATION:	PCR primer													
295	<400>	SEQUENCE:	4													
296	agtcaagctt	gtttatTTT	tccttactta	cagatgaagg											40	

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/08/961,083A

DATE: 06/30/2003
TIME: 11:37:27

Input Set : A:\340-SL-12-Dec-02.txt
Output Set: N:\CRF4\06272003\H961083A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 1368
Seq#:2; Xaa Pos. 456